

UNITED STATES DISTRICT COURT
NORTHERN DISTRICT OF CALIFORNIA

MARY L. SMITH, et al.,

Plaintiffs,

v.

GOOGLE, LLC,

Defendant.

Case No. 23-cv-03527-PCP (SVK)

DISCOVERY ORDER

Re: Dkt. No. 122, 123

Before the Court are several discovery disputes set forth in two Joint Submissions. Dkt. 122, 123. The Court determines that these issues may be resolved without oral argument. *See* Civ. L.R. 7-1(b).

I. DISCOVERY DISPUTES RE ALGORITHMS AND APPLICATIONS (Dkt. 122)

The Parties present three disputes in this Joint Submission, each of which is addressed below.

1. Blake Lemoine as Google Document Custodian.

Plaintiffs' request is **DENIED**. This Court is intimately familiar with the *Brown v. Google* litigation, including the technology at issue, the privacy issues arising from that technology and the vast technological challenges of discovery, including the identification and use of relevant custodians. Although some of the technical vocabulary in *Brown* may overlap with this case, and it may be necessary to employ some of the same discovery management tools, at a more granular level the cases present very differently. The Court disagrees that the Lemoine testimony relied upon by Plaintiffs in the Joint Submission demonstrates a sufficient nexus to this litigation. On the record before it, adding Mr. Lemoine as a custodian is simply not proportional to the needs of the case. Fed. R. Civ. Pro. 26(b)1.

////

1 **2. RFP 36 Regarding Algorithms, Artificial Intelligence and Automated**
 2 **Decision-Making Processes.**

3 Plaintiffs seek, via a proposed compromise, documents relating to “algorithms or AI
 4 systems used in the context of [Google Analytics], Firebase Analytics, Google Tag Manager, and
 5 any other tools implemented on the Subject Websites or applied to Website Visitor Data during
 6 the CLASS PERIOD....” Dkt. 122-1. In the Joint Submission, Plaintiffs explain that
 7 they are entitled to documents that demonstrate “how Google processes and monetizes Plaintiffs’
 8 and the Classes’ data after it has been intercepted, and the products and tools used by Google for
 9 these tasks.” Dkt. 122 at 6.

10 The Court agrees that Plaintiffs are entitled to discovery as to how Google “processes and
 11 monetizes” the subject data, however, how much discovery Plaintiffs are entitled to regarding the
 12 underlying products and tools used to “process and monetize” will be bounded by strictures of
 13 relevance and proportionality. Plaintiffs propose no such boundaries. Instead, relying again on
 14 the ill-fated testimony of Blake Lemoine, Plaintiffs appear to be seeking an extensive document
 15 production relating to Google’s AI products. Dkt. 122 at 6. As framed, Plaintiffs’ request is
 16 overly broad and not proportional to the needs of the litigation and is therefore **DENIED**.

17 Google states unequivocally that it has “agreed to produce documents demonstrating how
 18 data Google received from Google Analytics from the websites at issue may be used, and does not
 19 intend to withhold responsive and relevant non-privileged documents reflecting how Google may
 20 use GA data, including any ‘algorithms,’[sic] machine learning, or ads modeling.” Dkt. 122-1.
 21 This position, is, at a minimum, an appropriate place to start and comports with the Court’s ruling
 22 as to the third dispute, below. Accordingly, the Court **ORDERS Google to complete this**
 23 **production no later than December 17, 2025.**

24 ////

25 ////

26 ////

27 ////

28 ////

3. **Plaintiffs’ RFPs Concerning Applications and Application Related Data**
(RFPs 3¹, 9, 12, 16, 27, 38 and 39)

In addition to the documents relating to Google’s collection of data via Google Analytics (“GA”) from the Tax Websites, the production of which the Court has ordered to be completed by December 17, 2025, Plaintiffs seek production of documents relating to Google’s collection of data from the Tax Preparer’s applications (“apps”), and third party apps, via Google tools such as Firebase Analytics (“GA4F”). Dkt. 122 at 8. In support of this request, Plaintiffs point to two references in the Complaint of Google collecting data from “apps”. *Id.* at 9; Dkt. 35, ¶¶ 23, 24. Google challenges the request, arguing that: (1) the Complaint is framed as alleged wiretapping of websites via the GA product, not wiretapping of apps via GA4F, (*see* Dkt. 35, ¶¶ 1, 2); (2) the named Plaintiffs allege only use of the subject websites, not apps, (*id.*, ¶¶ 5-14); (3) the Tax Preparers are identified as “services that had the tracking pixel deployed on their websites,” (*id.*, ¶ 36; and (4) the class allegations are all identified by use of websites, not apps, (*id.*, ¶ 48).

In the Court’s experience, and as illustrated by the case cited by Defendant, discovery in class actions often addresses the tension between specific products identified in a complaint and additional products which may, under the facts alleged, be within the parameters of relevance and proportionality mandated by Rule 26. Dkt. 122 at 9 (citing *Alcala v. Monsanto Co.*, 2014 WL 1266204, at *5 (N.D. Cal. Mar. 24, 2014) (documents “related to products [plaintiff] did *not* use are not relevant.”).² Here, Plaintiffs’ reference to Google collecting data from apps is solely in the context of background; the references to “apps” are generic and do not relate to the Plaintiffs, the Tax Preparers or the specific websites which are the foundation for Plaintiffs’ claims. *See* Dkt. 35, ¶¶ 23-24. Significantly, the named Plaintiffs allege only use of the **websites**, and the Complaint specifically focuses an “embedded pixel” added to “**website** code.” *Id.*, ¶ 25 (emphasis added); *see also id.*, ¶ 26 (“Every time a user visits a **webpage** [with the code], the tracking code will

¹ The Parties also have a broader dispute regarding RFP 3 arising from the request inclusion of “[A]ll raw data.” This dispute is the subject of Dkt. 123 and is addressed in Section II below.

² The case relied on by Plaintiffs, *Pasadena Oil & Gas Wyoming LLC v. Montana Oil Properties Inc.*, 320 Fed. App’x. 675, 677 (9th Cir. 2009), does not relate to class actions and is therefore inapposite.

collect [purportedly] pseudonymous information about how that user interacted with *the page*.” (Emphasis added)); *id.*, ¶ 28 (“User *website* interactions and data collected by Google Analytics’ pixel are transmitted in real time to Google....” (Emphasis added)). In sum, the Complaint is about pixels on websites, not apps. The Court **DENIES** Plaintiffs’ requests for discovery related to the Tax Preparers’ websites and GAF4 as overbroad and not proportional to the needs of the case. Fed. R. Civ. P. 26.

II. DISCOVERY DISPUTE RE PRODUCTION OF CLASS EVENT DATA (Dkt. 123)

Plaintiffs seek production of all raw data received from the subject websites via GA during the class period. The Parties understandably have been wrangling the parameters of this request for nearly a year, and a third party, TaxAct, LLC (“TaxAct”), one of the relevant Tax Preparers, has weighed in on the dispute as well. *See* Dkt. 124. Plaintiffs have proposed at least one useful limitation as to the types of data to be included as “Class Event Data.” Dkt. 123-1 at 1. Google, finally at the time of the Joint Submission, helpfully proposes production of “an anonymized sample of the Google Analytics data associated with the Tax Websites’ relevant Google Analytics properties.” *Id.* The Court has more than a little experience with the identification and production of database samples in the context of Google privacy litigation. Sampling is particularly appropriate where, as here, production of “all raw data” is neither plausible nor even desirable (though Plaintiffs may disagree). Identifying a meaningful sample presents its own set of challenges, but neither Plaintiffs’ nor Google’s counsel should be unfamiliar with the approach. Accordingly, the Court **ORDERS** as follows:

- Google will produce an anonymized sample of the raw data received from the Tax Websites via Google Analytics;
- **No later than November 21, 2025**, the Parties are to meet and confer, either in person or by video conference, as to the types of data to be included, the relevant time frame, the size of the sample³ and a date for completion of production; and

³ In the Court’s experience, Plaintiffs’ proposed sample size is often overly large, failing to address the practical issues of production and defeating the efficiencies of a sample. Similarly, Google’s proposed sample may be absurdly small, inadequate for meaningful extrapolation to class-wide issues. The Parties are strongly encouraged to skip this initial step, and move directly

- If the Parties are unable to reach agreement on a sampling protocol, they may submit the issues in dispute to the Court **no later than 12 p.m. on November 26, 2025**, in the following format:

- A single table setting forth the disputed parameters on the *x* axis and the Parties' respective proposals, in brief without argument, on the *y* axis, *e.g.*:

Disputed Parameter [<i>e.g.</i> , Sample Size]	Plaintiffs' Position	Google's Position	Court's Use
--	----------------------	-------------------	-------------

- The table should be preceded by a paragraph outlining the agreed upon parameters of the sampling protocol.
- The Court will hold a hearing by video conference on **Tuesday, December 2nd, 2025 at 10 a.m.**, to address the open issues.

In conclusion, the Court notes that TaxAct, LLC appears in this action as an Objector and as such will receive notice of this Order. The use of an anonymized sample should address TaxAct's concerns.

SO ORDERED.

Dated: November 17, 2025



SUSAN VAN KEULEN
United States Magistrate Judge

to realistic, sound proposals.